CTL EPITOPES

How to Use Part I:

Section 1: HIV CTL Epitope Tables, Maps, Variability Plots, and Alignments by Protein

This section summarizes HIV-specific CTL epitopes arranged sequentially according to their location on the genome, organized by protein. We attempted to make this section as comprehensive as possible, requiring that the epitope be contained within a region of 30 or so amino acids, but not that the precise boundaries be defined. The HLA specificity of the epitope may or may not have been defined. For a table of the best characterized CTL epitopes that have precisely defined boundaries of 8-10 amino acids, with known associated HLA molecules, please see the review by Christian Brander and Bruce Walker in Section IV: *The HLA-class I restricted CTL response in HIV-1 infection.*

TABLES: each CTL epitope has a five part basic entry:

- **Location:** The amino acid positions of the epitope boundaries and the reference sequence are listed as given in the primary publication. Frequently, these positions as published are imprecise, and do not truly correspond to the numbering of the sequence, but they provide a reasonable guide to the peptide's approximate location in the protein. Also, in many cases the reference sequence identification was not provided.
- **Epitope:** The amino acid sequence of the epitope of interest as defined in the reference, based on the reference strain used in the study defining the epitope. On rare occasions, when only the epitope location and not the actual epitope was specified in the original publication, and the sequences were numbered inaccurately by the primary authors, we may have misrepresented the epitope's amino acid sequence.
- **Antigen:** The immunogen that stimulated the CTL response.
- **Species(HLA):** The species responding and HLA specificity of the epitope.
- **Reference**

Following each entry for a given CTL epitope is a brief comment explaining the context of the study that defined the epitope.

MAP:

The location and HLA restriction elements of CTL epitopes are indicated on protein sequences of the WEAU clone 1.60. This map is meant to provide the relative location of epitopes on a given protein, but the WEAU sequence may not actually carry the epitope of interest, as it may vary relative to the sequence for which the epitope was defined. Epitope regions are numbered, and the numbering on this map is used for the variability plots and the epitope alignments.

WEAU was chosen as the reference clone because it is one of the best characterized sequences currently available; the sequence was graciously provided prior to publication by George Shaw, and the manuscript describing the clone and sequence is in preparation. The clone was obtained from a coculture of this patient's PBMC's, first with normal donor PHA-stimulated lymphocytes for 14 days, and then with the H9 T-cell line for another 14 days. The blood specimen was obtained 15 days after the onset of clinical symptoms of acute (primary) infection, and 35 days after a single sexual encounter (receptive anal intercourse) with a partner whose virus was proven phylogenetically to be responsible for the transmission event. The single nucleotide deletion in nef in the WEAU 1.60 clone is NOT present in the patient's uncultured PBMCs where instead there is a "T." Thus, in the clone WEAU 1.60 nef is disrupted, but in the patient, the virus contains an intact nef gene in 10 out of 10 clones analyzed by PCR sequencing. The patient from whom WEAU 1.60 was derived is identified a "Patient #1" in N Engl J Med 324:954-960, 1991 and as "WEAU 0575" in Science 259:1749-1754, 1993. WEAU 1.60 and the virus isolate from which it was derived are SI (syncytium-inducing) strains. The full-length WEAU 1.60 provirus has been sequenced in its entirety by two different laboratories (G. Shaw and L. Hood) with 100% concordance.

VARIABILITY PLOTS:

The Shannon entropy for each column in the alignment is plotted for each position in the alignment (Korber et al., J. Virol. 68:7467-7481 1994). The entropy for perfectly conserved positions is zero, and higher numbers indicate relatively greater variability. Alignments of available full length (across a given protein) M group HIV-1 sequences from the 1995 HIV-1 sequence database were used as a basis to calculate the entropy for each po-

sition. The entropy scores can not be compared between different proteins, because the alignments consisted of different sets of sequences, and thus are not equivalent. The protein variability plots are internally consistent for a given protein, however, and thus provide an indication of the level of variability found in the region of a given epitope. The numbering of the epitopes corresponds to the epitope map.

ALIGNMENTS:

Following the numbering of the epitopes in the epitope map, alignments were generated from the protein sequence alignments in the HIV-1 genetic sequence database. All epitopes are aligned to the subtype B consensus (the most common amino acid found in subtype B in each position), with the sequence used to defined the epitope indicated directly beneath the B consensus. We used the 1995 HIV-1 database alignments for this section, although as the databases were being completed concurrently, the sequence compendium may have some additional sequences that were not in place when this section of the immunology database was generated. The sequence database alignments were modified if there were multiple insertions made to maintain the alignment across the epitope, in an attempt to optimize the alignment across the epitope and minimize insertions and deletions. A dash indicates identity, and a period indicates an insertion made to maintain the alignment. Epitopes with stop codons, frameshifts, or partial sequence were deleted. The consensus sequences for a subtype may not exactly reflect the sequences shown, as the consensus sequences were generated prior to the deletion of the problematic epitope sequences.

Section 2: Table of HIV CTL Epitopes Sorted by HLA Restricting Elements

This section is a table of the epitopes included in Section 1 that have known HLA restricting elements, sorted by the restricting element. Anchor and auxiliary residues for HLA molecules are listed, and if anchor residues with appropriate spacing are evident in the epitope, they are indicated by being written bold and underlined. This table provides minimal information about the epitopes; for more information see the table where epitopes are organized by location.

Section 3: References